Mathematics

AIMS Mathematics, 10(12): 28954–28980.
DOI: 10.3934/math.20251274
Received: 30 September 2025
Revised: 21 November 2025
Accepted: 28 November 2025
Published: 10 December 2025

https://www.aimspress.com/journal/Math

Research article

Median augmented ranked set sampling for estimation of the population mean with applications to body health data

Mahmoud Zuhier Aldrabseh[1,*], Khudhayr A. Rashedi[2], Ali A. Atoom[3], Tariq S. Alshammari[2] and Nesreen M. AL-Olaimat[4]

[1] Department of Mathematics, Faculty of Science, Zarqa University, Zarqa 13110, Jordan
[2] Department of Mathematics, College of Science, University of Ha'il, Hail 55425, Saudi Arabia
[3] Department of Mathematics, Faculty of Science, Ajloun National University, Ajloun, Jordan
[4] Department of Applied Science, Ajloun College, Al-Balqa Applied University, Ajoun, Jordan

* **Correspondence:** Email: mah.darabseh@gmail.com; Tel: +962790331657.

Abstract: Ranked set sampling (RSS) is a sampling design that combines random sampling with the judgment of researchers through preliminary ranking. The current study introduced a new generalization of RSS, called median-augmented ranked set sampling (MARSS), designed to further reduce the measurement cost and lessen the influence of outliers in estimating the population mean. The proposed MARSS estimator was compared with both simple random sampling (SRS) and RSS estimators. Its exact relative precision and bias were evaluated for a range of symmetric and skewed distributions under perfect ranking. A simulation study was also conducted to assess its performance under imperfect ranking, when using concomitant variables, in the presence of outliers, and when considering ranking cost efficiency. The variance and robustness were also interpreted in topological space. The theoretical results showed that the MARSS estimator was unbiased for symmetric distributions and achieved less variance than both RSS and SRS in unimodal symmetric distributions. Overall, MARSS is more precise than SRS and surpassed RSS in most scenarios, though some bias was observed for skewed distributions. Importantly, MARSS demonstrated a greater robustness to outliers than either SRS or RSS. Finally, the new sampling design was illustrated through an application to body health data analysis.

Keywords: ranked set sampling; simple random sampling; estimation; perfect ranking; imperfect ranking; concomitant variables; ranking cost; topological space
Mathematics Subject Classification: 60E05, 62D05, 62F10

1. Introduction

In many practical applications, such as environmental, agricultural, biological, and ecological studies, the process of taking actual measurements is often more expensive and time-consuming than performing a preliminary ranking based on judgment, concomitant variables, or visual inspection [1]. For instance, in estimating the average height of trees in a forest, it is generally easier to rank the trees visually than to measure their exact heights. To address the challenge of measurement cost, ranked set sampling (RSS) was introduced by McIntyre [2] in an agricultural setting to estimate the mean of pasture yield more efficiently. Since then, numerous modifications of RSS have been proposed to reduce measurement costs and improve estimation precision.

Researchers have introduced many innovative RSS methods aimed at improving sampling efficiency and flexibility. Among these are the median ranked set sampling (MRSS), which uses the median ranks [3]; the moving extreme RSS, which emphasizes the extreme ranks while varying the set sizes [4]; and L-ranked set sampling (LRSS) as a generalization of RSS through excluding certain extremes and replacing them with their nearest ranks [5]. Other designs have also been proposed, such as the systematic design of RSS [6]; and the recent except extreme RSS, which excludes the extreme ranks from all sets [7]. In addition to these single-stage procedures, researchers have also proposed multistage modifications, including double RSS [8], multistage ranked set sampling [9], double LRSS [10], and double except extreme RSS [11]. Further contributions in this area can be found in the works of [12–14], among others. These developments demonstrate the ongoing efforts to adapt RSS to different practical and theoretical requirements.

The RSS method and its modifications have been applied in a wide range of statistical inference problems and practical applications. For example, RSS designs have been used in the estimation of the cumulative distribution function [15], median [16], in quality control estimation [17,18], and in the estimation of variance [19] and ratio [12]. Further applications include regression estimation [20], estimation of reliability models [21–23], estimation of shape and scale parameters [24], in small area estimation [25,26], and more recent work on parameter estimation under RSS variations [27,28]. These various applications highlight the precision of RSS as a sampling design and its importance in both theoretical and applied statistics.

The aim of this study is to propose and evaluate a new variation of RSS that can improve mean estimation by reducing the effect of outliers. The proposed design, called median-augmented RSS (MARSS), will be compared with other traditional sampling methods. MARSS is expected to improve the efficiency and accuracy of mean estimation, offer more robust sampling against outliers, and thus reduce the time and cost of data collection. The importance of this study lies in its introduction of a new sampling method. To the best of our knowledge, this is the first study to explore the connection between topological spaces and ordered sampling methods, providing a novel framework for understanding and improving sampling efficiency. It can also be applied in estimating the other parameters of probability distributions under different truncation combinations, as in [29], reliability models, quality control, and other population characteristics, in addition to the practical application in data collection.

The remainder of the paper is structured as follows: Section 2 presents the foundation of the existing sampling method; Section 3 suggests and explains the proposed sampling design; Section 4 gives a comparative study of the new estimator under imperfect ranking with some estimators and a discussion of other cases, such as ranking using the concomitant variable, the impact of outliers, and ranking cost; Section 5 interprets the variance and robustness aspects in topological space; Section 6 presents applications and a numerical example; and finally, Section 7 gives the conclusions and future directions.

2. Existing sampling designs

This section presents a brief overview of existing sampling methods, highlighting the estimators of the population mean and measures of precision.

2.1. Simple random sampling

Simple random sampling (SRS) is a widely used probability sampling. Let $X_1, X_2, \ldots, X_n$ be a random sample from a population with a probability density function (PDF) $f(x)$ and cumulative density function (CDF) $F(x)$. The SRS estimator of the population mean μ is given by:

$$\hat{\mu}_{SRS} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i. \quad (1)$$

The variance of that estimator is:

$$V(\bar{X}) = \frac{\sigma^2}{n}, \quad (2)$$

where σ^2 is the population variance.

2.2. Ranked set sampling

RSS was introduced by McIntyre [2] to estimate the average pasture yield using preliminary judgment ranking. Independently, Takahasi and Wakimoto [30] established the mathematical foundation of RSS. The procedure can be summarized as follows:

(1) Randomly select m^2 units from the population and divide them into *m* sets.
(2) Rank the unit within each set using judgment, visual inspection, or an auxiliary variable related to the study variable.
(3) From the i^{th} set, measure the i^{th} ranked unit, for $i = 1, 2, \ldots, m$.
(4) Repeat Steps 1–3 over *c* cycles to obtain a final sample of size *n* = *mc*.

Let $X_{(i)}^{j}$ denote the i^{th} ranked unit in the j^{th} cycle. Then the RSS estimator of μ under perfect ranking is given by:

$$\hat{\mu}_{RSS} = \frac{1}{mc}\sum_{j=1}^{c}\sum_{i}^{m} X_{(i)}^{j}. \quad (3)$$

Here, $\hat{\mu}_{RSS}$ is an unbiased estimator [31]. The variance of the RSS estimator $\hat{\mu}_{RSS}$ is

$$V(\hat{\mu}_{RSS}) = \frac{1}{m^2 c}\sum_{i=1}^{m} \sigma_{(i)}^2, \quad (4)$$

where $\sigma_{(i)}^2$ is the variance of the i^{th} ordered statistic. It can be evaluated from its PDF as:

$$f_{(i)}(x) = \frac{m!}{(i-1)!(m-i)!} f(x)[F(x)]^{i-1}[1 - F(x)]^{m-i}. \quad (5)$$

For more details on ordered statistic and their moments, see [32]. The relative precision (RP) of the RSS estimator compared to SRS is:

$$RP = \frac{1}{1 - \frac{1}{m\sigma^2}\sum_{i=1}^{m} \tau_{(i)}^2}, \tag{6}$$

where $\tau_{(i)} = \mu_{(i)} - \mu$ and $\mu_{(i)}$ represents the expected value of the i^{th} ordered statistic.

3. Proposed sampling design (MARSS)

In this study, we propose a new modification of ranked set sampling (RSS) designed to provide greater robustness in the presence of outliers. The new design is termed the median-augmented ranked set sampling (MARSS). The procedure for implementing MARSS can be outlined as follows:

(1) Randomly select m samples of each size m.
(2) Rank the elements within each set using any costless procedure.
(3) Define a coefficient $h = [\delta m]$, where $0 \leq \delta < 0.5$, and $[v]$ is the largest integer less than or equal to v.
(4) If m is odd, measure the $((m+1)/2)^{th}$ ranked elements in the first h sets and last h sets, which means for $i = 1, \cdots, h$ and $i = m - h + 1, \cdots m$. In the remaining sets, measure the $(i)^{th}$ elements in the i^{th} set, for $i = h + 1, \cdots, m - h$.
(5) If m is even, measure the $((m)/2)^{th}$ ranked elements in the first h sets at $i = 1, \cdots, h$, and $((m+2)/2)^{th}$ in the last h sets at $i = m - h + 1, \cdots m$. In the remaining sets, measure the $(i)^{th}$ elements in the i^{th} set, for $i = h + 1, \cdots, m - h$.
(6) Steps 1–5 can be repeated c cycles to obtain a larger sample $n = cm$.

For more clarification and without loss of generality, suppose that the number of cycles is one ($c = 1$). The following example includes some cases.

Example 3.1. To demonstrate the MARSS procedure, we consider two cases with different set sizes and coefficients.

Case 1. At the coefficient *h = 1* and the odd set size *m = 5*.

$$\underbrace{\begin{Bmatrix} X_{11} & X_{12} & X_{13} & X_{14} & X_{15} \\ X_{21} & X_{22} & X_{23} & X_{24} & X_{25} \\ X_{31} & X_{32} & X_{33} & X_{34} & X_{35} \\ X_{41} & X_{42} & X_{43} & X_{44} & X_{45} \\ X_{51} & X_{52} & X_{53} & X_{54} & X_{55} \end{Bmatrix}}_{Selecting\ m\ random\ samples} \Rightarrow \underbrace{\begin{Bmatrix} X_{1(1:5)} & X_{1(2:5)} & \boldsymbol{X_{1(3:5)}} & X_{1(4:5)} & X_{1(5:5)} \\ X_{2(1:5)} & \boldsymbol{X_{2(2:5)}} & X_{2(3:5)} & X_{2(4:5)} & X_{2(5:5)} \\ X_{3(1:5)} & X_{3(2:5)} & \boldsymbol{X_{3(3:5)}} & X_{3(4:5)} & X_{3(5:5)} \\ X_{4(1:5)} & X_{4(2:5)} & X_{4(3:5)} & \boldsymbol{X_{4(4:5)}} & X_{4(5:5)} \\ X_{5(1:5)} & X_{5(2:5)} & \boldsymbol{X_{5(3:5)}} & X_{5(4:5)} & X_{5(5:5)} \end{Bmatrix}}_{Ranking\ and\ selecting\ for\ actual\ measurement}$$

$$\Rightarrow \underbrace{\begin{Bmatrix} X_{1(3:5)} \\ X_{2(2:5)} \\ X_{3(3:5)} \\ X_{4(4:5)} \\ X_{5(3:5)} \end{Bmatrix}}_{The\ selected\ elements}.$$

Case 2. At the coefficient *h = 2* and the even set size *m = 6*.

$$\underbrace{\begin{Bmatrix} X_{11} & X_{12} & X_{13} & X_{14} & X_{15} & X_{16} \\ X_{21} & X_{22} & X_{23} & X_{24} & X_{25} & X_{26} \\ X_{31} & X_{32} & X_{33} & X_{34} & X_{35} & X_{36} \\ X_{41} & X_{42} & X_{43} & X_{44} & X_{45} & X_{46} \\ X_{51} & X_{52} & X_{53} & X_{54} & X_{55} & X_{56} \\ X_{61} & X_{62} & X_{63} & X_{64} & X_{65} & X_{66} \end{Bmatrix}}_{Selecting\ m\ random\ samples} \Rightarrow \underbrace{\begin{Bmatrix} X_{1(1:6)} & X_{1(2:6)} & \boldsymbol{X_{1(3:6)}} & X_{1(4:6)} & X_{1(5:6)} & X_{1(6:6)} \\ X_{2(1:6)} & X_{2(2:6)} & \boldsymbol{X_{2(3:6)}} & X_{2(4:6)} & X_{2(5:6)} & X_{2(6:6)} \\ X_{3(1:6)} & X_{3(2:6)} & \boldsymbol{X_{3(3:6)}} & X_{3(4:6)} & X_{3(5:6)} & X_{3(6:6)} \\ X_{4(1:6)} & X_{4(2:6)} & X_{4(3:6)} & \boldsymbol{X_{4(4:6)}} & X_{4(5:6)} & X_{4(6:6)} \\ X_{5(1:6)} & X_{5(2:6)} & X_{5(3:6)} & \boldsymbol{X_{5(4:6)}} & X_{5(5:6)} & X_{5(6:6)} \\ X_{6(1:6)} & X_{6(2:6)} & X_{6(3:6)} & \boldsymbol{X_{6(4:6)}} & X_{6(5:6)} & X_{6(6:6)} \end{Bmatrix}}_{Ranking\ and\ selecting\ for\ actual\ measurement}$$

$$\Rightarrow \underbrace{\begin{Bmatrix} X_{1(3:6)} \\ X_{2(3:6)} \\ X_{3(3:6)} \\ X_{4(4:6)} \\ X_{5(4:6)} \\ X_{6(4:6)} \end{Bmatrix}}_{The\ selected\ elements} .$$

Further cases of the MARSS design are illustrated in Figure 1. Some well-known schemes that arise as special cases of the MARSS. For example, when $h = 0$, the MARSS design reduces to the RSS design. Moreover, if m is odd and $(m-1)/2 = h$, or if m is even and $(m-2)/2 = h,$ the MARSS design coincides with MRSS. In certain cases, the MARSS design also aligns with LRSS.

The estimator of the population mean at m set sizes, odd and even respectively, can be defined as:

$$\hat{\mu}_{MARSS,Odd} = \frac{1}{mc}\sum_{j=1}^{c}\left[\sum_{i=1}^{h} X^{j}_{i\left(\frac{m+1}{2}\right)} + \sum_{i=h+1}^{m-h} X^{j}_{i(i)} + \sum_{i=m-h+1}^{m} X^{j}_{i\left(\frac{m+1}{2}\right)}\right], \tag{7}$$

and

$$\hat{\mu}_{MARSS,Even} = \frac{1}{mc}\sum_{j=1}^{c}\left[\sum_{i=1}^{h} X^{j}_{i\left(\frac{m}{2}\right)} + \sum_{i=h+1}^{m-h} X^{j}_{i(i)} + \sum_{i=m-h+1}^{m} X^{j}_{i\left(\frac{m}{2}+1\right)}\right]. \tag{8}$$

The expectations of the corresponding estimator are:

$$E\left(\hat{\mu}_{MARSS,Odd}\right) = \frac{1}{m}\left[2h\mu_{\left(\frac{m+1}{2}\right)} + \sum_{i=h+1}^{m-h} \mu_{(i)}\right], \tag{9}$$

and

$$E\left(\hat{\mu}_{MARSS,Even}\right) = \frac{1}{m}\left[h\mu_{\left(\frac{m}{2}\right)} + h\mu_{\left(\frac{m}{2}+1\right)} + \sum_{i=h+1}^{m-h} \mu_{(i)}\right]. \tag{10}$$

The corresponding variances are given by:

$$\begin{aligned} V\left(\hat{\mu}_{MARSS,Odd}\right) &= V\left(\frac{1}{mc}\sum_{j=1}^{c}\left[\sum_{i=1}^{h} X^{j}_{i\left(\frac{m+1}{2}\right)} + \sum_{i=h+1}^{m-h} X^{j}_{i(i)} + \sum_{i=m-h+1}^{m} X^{j}_{i\left(\frac{m+1}{2}\right)}\right]\right) \\ &= \frac{1}{m^2 c}\left[2h\sigma^2_{\left(\frac{m+1}{2}\right)} + \sum_{i=h+1}^{m-h} \sigma^2_{(i)}\right], \end{aligned} \tag{11}$$

and

$$V(\hat{\mu}_{MARSS,Even}) = V\left(\frac{1}{mc}\sum_{j=1}^{c}\left[\sum_{i=1}^{h} X^{j}_{i\left(\frac{m}{2}\right)} + \sum_{i=h+1}^{m-h} X^{j}_{i(i)} + \sum_{i=m-h+1}^{m} X^{j}_{i\left(\frac{m}{2}+1\right)}\right]\right)$$
$$= \frac{1}{m^2 c}\left[h\sigma^2_{\left(\frac{m}{2}\right)} + h\sigma^2_{\left(\frac{m}{2}+1\right)} + \sum_{i=h+1}^{m-h}\sigma^2_{(i)}\right]. \quad (12)$$

a) At *h = 1* and *m = 3*

$X_{1(1:3)}$	$\boldsymbol{X_{1(2:3)}}$	$X_{1(3:3)}$
$X_{2(1:3)}$	$\boldsymbol{X_{2(2:3)}}$	$X_{2(3:3)}$
$X_{3(1:3)}$	$\boldsymbol{X_{3(2:3)}}$	$X_{3(3:3)}$

b) At *h = 1* and *m = 6*

$X_{1(1:6)}$	$X_{1(2:6)}$	$\boldsymbol{X_{1(3:6)}}$	$X_{1(4:6)}$	$X_{1(5:6)}$	$X_{1(6:6)}$
$X_{2(1:6)}$	$\boldsymbol{X_{2(2:6)}}$	$X_{2(3:6)}$	$X_{2(4:6)}$	$X_{2(5:6)}$	$X_{2(6:6)}$
$X_{3(1:6)}$	$X_{3(2:6)}$	$\boldsymbol{X_{3(3:6)}}$	$X_{3(4:6)}$	$X_{3(5:6)}$	$X_{3(6:6)}$
$X_{4(1:6)}$	$X_{4(2:6)}$	$X_{4(3:6)}$	$\boldsymbol{X_{4(4:6)}}$	$X_{4(5:6)}$	$X_{4(6:6)}$
$X_{5(1:6)}$	$X_{5(2:6)}$	$X_{5(3:6)}$	$X_{5(4:6)}$	$\boldsymbol{X_{5(5:6)}}$	$X_{5(6:6)}$
$X_{6(1:6)}$	$X_{6(2:6)}$	$X_{6(3:6)}$	$\boldsymbol{X_{6(4:6)}}$	$X_{6(5:6)}$	$X_{6(6:6)}$

c) At *h = 1* and *m = 4*

$X_{1(1:4)}$	$\boldsymbol{X_{1(2:4)}}$	$X_{1(3:4)}$	$X_{1(4:4)}$
$X_{2(1:4)}$	$\boldsymbol{X_{2(2:4)}}$	$X_{2(3:4)}$	$X_{2(4:4)}$
$X_{3(1:4)}$	$X_{3(2:4)}$	$\boldsymbol{X_{3(3:4)}}$	$X_{3(4:4)}$
$X_{4(1:4)}$	$X_{4(2:4)}$	$\boldsymbol{X_{4(3:4)}}$	$X_{4(4:4)}$

d) At *h = 2* and *m = 5*

$X_{1(1:5)}$	$X_{1(2:5)}$	$\boldsymbol{X_{1(3:5)}}$	$X_{1(4:5)}$	$X_{1(5:5)}$
$X_{2(1:5)}$	$X_{2(2:5)}$	$\boldsymbol{X_{2(3:5)}}$	$X_{2(4:5)}$	$X_{2(5:5)}$
$X_{3(1:5)}$	$X_{3(2:5)}$	$\boldsymbol{X_{3(3:5)}}$	$X_{3(4:5)}$	$X_{3(5:5)}$
$X_{4(1:5)}$	$X_{4(2:5)}$	$\boldsymbol{X_{4(3:5)}}$	$X_{4(4:5)}$	$X_{4(5:5)}$
$X_{5(1:5)}$	$X_{5(2:5)}$	$\boldsymbol{X_{5(3:5)}}$	$X_{5(4:5)}$	$X_{5(5:5)}$

Figure 1. Selected MARSS scenarios.

Theorem 3.1. Assume the parent distribution $f(x)$ is symmetric, and then the estimator $\hat{\mu}_{MARSS}$ is an unbiased estimator of μ.

Proof. It is known that $\sum_{i=1}^{m}\mu_{(i)} = m\mu$. Moreover, by the symmetry property [32], we have

$$\mu - \mu_{(i)} = \mu_{(m-i+1)} - \mu.$$

For odd m:

$$E(\hat{\mu}_{MARSS,Odd}) = \frac{1}{m}\left[2h\mu_{\left(\frac{m+1}{2}\right)} + \sum_{i=h+1}^{m-h}\mu_{(i)}\right] = \frac{1}{m}[2h\mu + (m-2h)\mu] = \mu.$$

For even m:

$$E(\hat{\mu}_{MARSS,Even}) = \frac{1}{m}\left[h\mu_{\left(\frac{m}{2}\right)} + h\mu_{\left(\frac{m}{2}+1\right)} + \sum_{i=h+1}^{m-h}\mu_{(i)}\right] = \frac{1}{m}[2h\mu + (m-2h)\mu] = \mu.$$

The proof is complete.

Theorem 3.2. Assume the parent distribution $f(x)$ is symmetric and unimodal. Then

a) $V(\hat{\mu}_{MARSS}) \le V(\hat{\mu}_{RSS})$.

b) $V(\hat{\mu}_{MARSS}) \le V(\hat{\mu}_{SRS})$.

Proof.

a) For a symmetric unimodal distribution, the variances of the order statistics decrease as the ranks move toward the median

$$V\left(\hat{\mu}_{MARSS,Odd}\right)=\frac{1}{m^2c}\left[h\sigma^2_{\left(\frac{m+1}{2}\right)}+\sum_{i=h+1}^{m-h}\sigma^2_{(i)}+h\sigma^2_{\left(\frac{m+1}{2}\right)}\right],$$

while

$$V(\hat{\mu}_{RSS})=\frac{1}{m^2c}\sum_{i=1}^{m}\sigma^2_{(i)}=\frac{1}{m^2c}\left[\sum_{i=1}^{h}\sigma^2_{(i)}+\sum_{i=h+1}^{m-h}\sigma^2_{(i)}+\sum_{i=m-h+1}^{m}\sigma^2_{(i)}\right].$$

As mentioned earlier, because the density is higher as we move toward the median ranks, the median order statistic has the smallest variance [5,32], hence

$$h\sigma^2_{\left(\frac{m+1}{2}\right)}\le\sum_{i=1}^{h}\sigma^2_{(i)}\,,h\sigma^2_{\left(\frac{m+1}{2}\right)}\le\sum_{i=m-h+1}^{m}\sigma^2_{(i)}.$$

Thus, $V\left(\hat{\mu}_{MARSS,Odd}\right)\le V(\hat{\mu}_{RSS})$. The argument extends directly to the even m case with variance:

$$V\left(\hat{\mu}_{MARSS,Even}\right)=\frac{1}{m^2c}\left\{k\sigma^2_{\left(\frac{m}{2}\right)}+\sum_{i=h+1}^{m-h}\sigma^2_{(i)}+k\sigma^2_{\left(\frac{m}{2}+1\right)}\right\}.$$

Finally, the inequality with respect to SRS follows from the decomposition

$$\sum_{i=1}^{m}\sigma^2_{(i)}=m\sigma^2-\sum_{i=1}^{m}\left(\mu_{(i)}-\mu\right)^2$$

and the established results show that $V(\hat{\mu}_{RSS})\le V(\hat{\mu}_{SRS})$ [33]. Since

$$V(\hat{\mu}_{MARSS})\le V(\hat{\mu}_{RSS}),$$

we also have

$$V(\hat{\mu}_{MARSS})\le V(\hat{\mu}_{SRS}).$$

The proof is now complete.

In the case of a skewed distribution, $\hat{\mu}_{MARSS}$ is no longer unbiased. The bias (B) is given by:

$$B(\hat{\mu}_{MARSS})=E(\hat{\mu}_{MARSS})-\mu. \tag{13}$$

Here, $E(\hat{\mu}_{MARSS})$ is given in Eq (9) to Eq (10). Therefore, by using Eq (11) to Eq (12), the mean squared error (MSE) of $\hat{\mu}_{MARSS}$ is then:

$$MSE(\hat{\mu}_{MARSS})=V(\hat{\mu}_{MARSS})+[B(\hat{\mu}_{MARSS})]^2. \tag{14}$$

Accordingly, the RP of $\hat{\mu}_{MARSS}$ compared to the SRS estimator can be expressed for odd and even *m* as:

$$RP_{Odd}=\frac{V(\hat{\mu}_{SRS})}{MSE(\hat{\mu}_{MARSS.Odd})}=\frac{m\sigma^2}{\left\{2h\sigma^2_{\left(\frac{m+1}{2}\right)}+\sum_{i=h+1}^{m-h}\sigma^2_{(i)}\right\}+c\left\{2h\mu_{\left(\frac{m+1}{2}\right)}+\sum_{i=h+1}^{m-h}\mu_{(i)}-m\mu\right\}^2}, \tag{15}$$

$$RP_{Even}=\frac{V(\hat{\mu}_{SRS})}{MSE\left(\hat{\mu}_{MARSS,Even}\right)}$$

$$=\frac{m\sigma^2}{\left\{h\sigma^2_{\left(\frac{m}{2}\right)}+\sum_{i=h+1}^{m-h}\sigma^2_{(i)}+h\sigma^2_{\left(\frac{m}{2}+1\right)}\right\}+c\left\{h\mu_{\left(\frac{m}{2}\right)}+\sum_{i=h+1}^{m-h}\mu_{(i)}+h\mu_{\left(\frac{m}{2}+1\right)}-m\mu\right\}^2}. \tag{16}$$

Once the moments of the order statistics used in the MARSS design are derived from the PDF in Eq (5), the corresponding performance measures (RP and B) can be calculated directly.

4. Results and discussion

In this section, we present a comparative analysis of the MARSS estimator against several existing sampling methods (SRS and RSS). The evaluation covers a range of scenarios, including perfect for some distributions and imperfect ranking, the use of concomitant variables, the impact of outliers, and cost efficiency.

4.1. Perfect ranking

To test the performance of the MARSS estimator, using Walframe Mathematica 13, we evaluate the analytical values of RP and bias of the estimators under RSS and MARSS with respect to the corresponding SRS estimator, for the scenarios of $h = 0,1$, and 2, and the scenarios of $m = 3$ to 8. The RP evaluated exactly, using Mathematica 13.0, for a group of symmetrical distributions: *Uniform(0,1), N(0,1), Logistic(2,1), StudentT(5),* and *Beta(5,5).* In addition to RP and B of the estimators for the skewed distributions: *Beta(3,5), Gamma(4,1), Weibull(1,4), Exp(1), Chi(3), LogNormal(0,1),* and *HalfNormal(5)*.

Table 1 reports the RPs of the RSS and MARSS estimators compared to the SRS for symmetric distributions. In all cases, RP is greater than one, meaning that both RSS and MARSS are more efficient than SRS. As m increases, RP also increases. Additionally, RPs vary depending on the distribution. For all distributions except the uniform case, MARSS with $h = 2$ is the most efficient, while MARSS with $h = 1$ is more efficient than RSS (Figure 2), which supports the findings in Theorem 2.

Table 2 illustrates the RP for some skewed distributions. In all considered cases, RP is greater than one, indicating that both RSS and MARSS estimators are more efficient than the SRS estimator. As m increases, the RP increases for RSS and for MARSS in some cases of m. Also, the RPs vary from one distribution to another. In most cases, especially at a small set size m ($m = 3$ or 4) and some of the larger m, the MARSS estimator with $h = 1$ is the most efficient (Figure 2). The superiority of RSS in some cases is due to the bias of MARSS in certain skewed distributions.

It is known that the RSS and SRS estimators are unbiased estimators. For this reason, Table 3 presents only the bias of the MARSS estimator with $h = 1$ and $h = 2$. The results show that the bias of the MARSS estimators varies significantly from one distribution to another; it is negligible in some cases, such as Beta and Weibull, and large in the remaining cases. In general, the bias of MARSS with $h = 1$ is smaller than MARSS with $h = 2$.

Table 1. RP of the mean estimators for selected symmetrical distributions under RSS and MARSS.

Distribution	m	RSS ($h=0$)	MARSS ($h=1$)	MARSS ($h=2$)
$Uniform(0,1)$	3	**2.00**	1.67	**
	4	**2.50**	2.08	**
	5	**3.00**	2.44	2.33
	6	**3.50**	2.88	2.72
	7	**4.00**	3.29	3.17
	8	**4.50**	3.75	3.46
$N(0,1)$	3	1.91	**2.23**	**
	4	2.35	**2.77**	**
	5	2.77	3.37	**3.49**
	6	3.19	3.89	**4.06**
	7	3.59	4.42	**4.69**
	8	4.00	4.91	**5.25**
$Logistic(2,1)$	3	1.84	**2.55**	**
	4	2.22	**3.16**	**
	5	2.58	3.89	**4.17**
	6	2.93	4.44	**4.85**
	7	3.27	5.02	**5.66**
	8	3.60	5.52	**6.29**
$StudentT(5)$	3	1.76	**2.86**	**
	4	2.08	**3.54**	**
	5	2.38	4.38	**4.76**
	6	2.66	4.96	**5.54**
	7	2.93	5.59	**6.48**
	8	3.19	6.12	**7.20**
$Beta(5,5)$	3	1.95	2.07	**
	4	2.41	2.57	**
	5	2.87	3.11	**3.15**
	6	3.32	3.60	**3.68**
	7	3.77	4.10	**4.22**
	8	4.21	4.59	**4.73**

Table 2. RP of the mean estimators for selected skewed distributions under RSS and MARSS.

Distribution	m	RSS	$h=1$	$h=2$
$Beta(3{,}5)$	3	1.94	**2.03**	**
	4	2.40	**2.52**	**
	5	2.86	**3.00**	**3.01**
	6	3.31	**3.46**	**3.46**
	7	3.75	**3.91**	3.89
	8	4.20	**4.34**	4.30
$Gamma(4{,}1)$	3	1.83	**2.22**	**
	4	2.21	**2.66**	**
	5	2.58	**3.02**	3.01
	6	2.94	3.33	**3.46**
	7	3.29	3.58	**3.89**
	8	3.64	3.82	**4.30**
$Weibull(1{,}4)$	3	1.93	**2.15**	**
	4	2.38	**2.67**	**
	5	2.82	3.23	**3.31**
	6	3.25	3.73	**3.85**
	7	3.68	4.25	**4.43**
	8	4.11	4.73	**4.95**
$Exp(1)$	3	1.64	**2.25**	**
	4	1.92	**2.44**	**
	5	2.19	**2.44**	2.23
	6	2.45	**2.49**	2.14
	7	**2.70**	2.46	1.91
	8	**2.94**	2.50	1.84
$Chi(3)$	3	1.71	**2.23**	**
	4	2.03	**2.51**	**
	5	2.34	**2.64**	2.47
	6	2.63	**2.77**	2.46
	7	**2.92**	2.81	2.29
	8	**3.20**	2.90	2.25
$LogNormal(0{,}1)$	3	1.34	**3.41**	**
	4	1.47	**3.35**	**
	5	1.59	**3.08**	2.80
	6	1.70	**2.94**	2.51
	7	1.80	**2.75**	2.14
	8	1.89	**2.67**	1.98
$HalfNormal(5)$	3	1.84	**2.01**	**
	4	2.24	**2.37**	**
	5	**2.63**	2.62	2.49
	6	**3.01**	2.87	2.63
	7	**3.39**	3.05	2.61
	8	**3.76**	3.25	2.67

Table 3. Bias of the mean estimators for selected skewed distributions under MARSS.

Distribution	m	MARSS (*h* = *1*)	MARSS (*h* = *2*)
$Beta(3,5)$	3	**-0.01**	**
	4	**-0.01**	**
	5	**-0.01**	**-0.01**
	6	**-0.01**	**-0.01**
	7	**-0.01**	**-0.01**
	8	**-0.01**	**-0.01**
$Gamma\ (4,1)$	3	-0.18	**
	4	-0.18	**
	5	-0.20	-0.23
	6	-0.19	-0.23
	7	-0.19	-0.25
	8	**-0.18**	-0.24
$Weibull\ (1,4)$	3	**0.00**	**
	4	**0.00**	**
	5	**0.00**	**0.00**
	6	**0.00**	**0.00**
	7	**0.00**	**0.00**
	8	**0.00**	**0.00**
$Exp(1)$	3	**-0.17**	**
	4	**-0.17**	**
	5	**-0.18**	-0.22
	6	**-0.18**	-0.22
	7	**-0.17**	-0.23
	8	**-0.16**	-0.22
$Chi(3)$	3	**-0.35**	**
	4	**-0.35**	**
	5	**-0.38**	-0.45
	6	**-0.36**	-0.45
	7	**-0.36**	-0.47
	8	**-0.34**	-0.46
$LogNormal(0,1)$	3	**-0.40**	**
	4	**-0.40**	**
	5	**-0.43**	-0.49
	6	**-0.41**	-0.49
	7	**-0.41**	-0.52
	8	**-0.39**	-0.51
$HalfNormal(5)$	3	**-0.33**	**
	4	**-0.33**	**
	5	**-0.36**	-0.43
	6	**-0.34**	-0.43
	7	**-0.34**	-0.45
	8	**-0.33**	-0.44

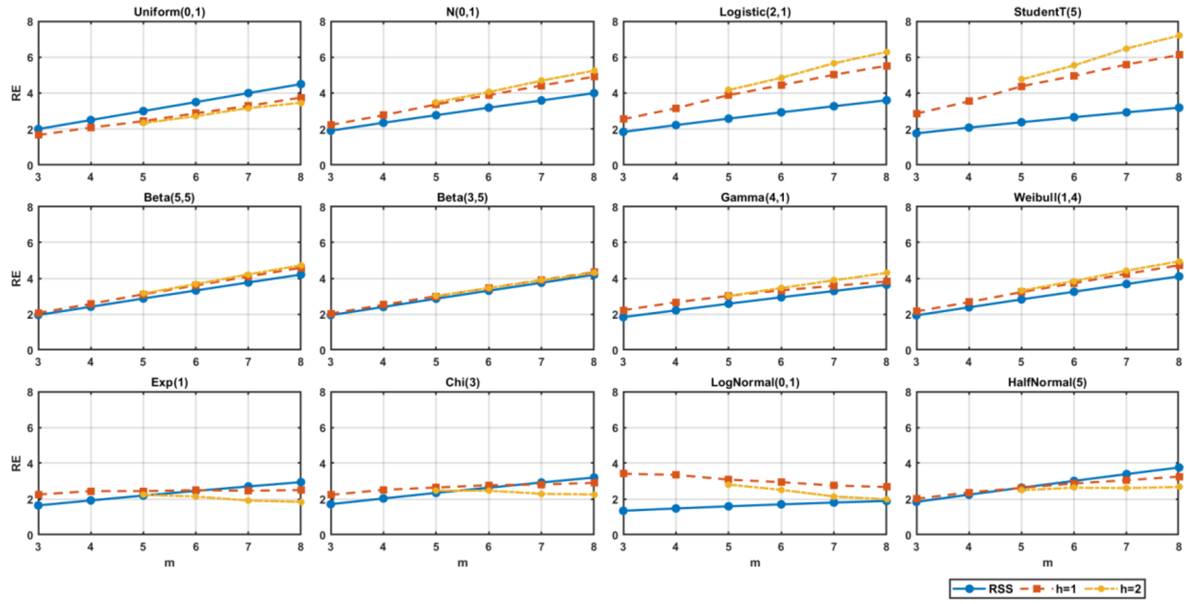

Figure 2. RP of MARSS and RSS mean estimators under perfect ranking.

In summary, the results of RP and bias demonstrate that the MARSS (h = 2) design is the most efficient sampling design for symmetric unimodal distributions, achieving the highest RP values. For skewed distributions, MARSS (h = 1) generally is the best choice, particularly when the set size is small. However, this design exhibits bias in some of the skewed distributions.

4.2. Imperfect ranking

In the previous subsection, we assumed perfect ranking, meaning that the distribution of ranked units is identical to the distribution of the corresponding ordered statistics. In practice, however, the assumption of perfect ranking in RSS schemes is often unrealistic, as judgment errors can occur. This subsection examines the impact of ranking errors on the performance of the proposed MARSS design. The case of imperfect ranking was first formalized by Dell and Clutter [32] using a single-valued function:

$$Y_{(i)} = g\left(X_{[i]}, e_i\right),$$

where e_i represents the ranking error for the i^{th} ranked unit, $Y_{(i)}$ denotes the variable used for ranking (subject to error), while the true measurement is taken from $X_{[i]}^{*}$ (with ranking error). The simplest noisy model is

$$Y_{(i)} = X_{[i]}^{*} + e_i, \tag{17}$$

where $X_{[1]}^{*}, X_{[2]}^{*}, \dots, X_{[m]}^{*}$ and $e_1, e_2, \cdots, e_m$ are mutually independent, and the errors are assumed to follow a normal distribution,

$$e_i \sim N(0, \sigma_e^2).$$

The RSS estimator of μ under this error model is

$$\hat{\mu}_{RSS,Imp} = \frac{1}{mc} \sum_{j=1}^{c} \sum_{i=1}^{m} X_{[i]}^{*(j)}. \tag{18}$$

Similarly, the MARSS estimator $\hat{\mu}_{MARSS,Imp}$ under an imperfect ranking is given by:

For odd *m*:

$$\hat{\mu}_{MARSS,Odd,Imp} = \frac{1}{mc}\sum_{j=1}^{c}\left[\sum_{i=1}^{h} X^{*(j)}_{\left[\frac{m+1}{2}\right]} + \sum_{i=h+1}^{m-h} X^{*(j)}_{[i]} + \sum_{i=m-h+1}^{m} X^{*(j)}_{\left[\frac{m+1}{2}\right]}\right]. \tag{19}$$

For even *m*,

$$\hat{\mu}_{MARSS,Even,Imp} = \frac{1}{mc}\sum_{j=1}^{c}\left[\sum_{i=1}^{h} X^{*(j)}_{\left[\frac{m}{2}\right]} + \sum_{i=h+1}^{m-h} X^{*(j)}_{[i]} + \sum_{i=m-h+1}^{m} X^{*(j)}_{\left[\frac{m}{2}+1\right]}\right]. \tag{20}$$

The corresponding variances can be derived as:

$$V\left(\hat{\mu}_{MARSS,Odd,Imp}\right) = \frac{1}{m^2c}\left[\sum_{i=1}^{h} \sigma^2_{X_{\left[\frac{m+1}{2}\right]}} + \sum_{i=h+1}^{m-h} \sigma^2_{X_{[i]}} + \sum_{i=m-h+1}^{m} \sigma^2_{X_{\left[\frac{m}{2}+1\right]}}\right]$$
$$= \frac{\sigma_e^2}{mc} + \frac{1}{m^2c}\left[2h\sigma^2_{Y_{\left(\frac{m+1}{2}\right)}} + \sum_{i=h+1}^{m-h} \sigma^2_{Y_{(i)}}\right] \tag{21}$$

and

$$V\left(\hat{\mu}_{MARSS,Even,Imp}\right) = \frac{1}{m^2c}\left[\sum_{i=1}^{h} \sigma^2_{X_{\left[\frac{m}{2}\right]}} + \sum_{i=h+1}^{m-h} \sigma^2_{X_{[i]}} + \sum_{i=m-h+1}^{m} \sigma^2_{X_{\left[\frac{m}{2}+1\right]}}\right]$$
$$= \frac{\sigma_e^2}{mc} + \frac{1}{m^2c}\left[h\sigma^2_{Y_{\left(\frac{m}{2}\right)}} + h\sigma^2_{Y_{\left(\frac{m}{2}+1\right)}} + \sum_{i=h+1}^{m-h} \sigma^2_{Y_{(i)}}\right]. \tag{22}$$

It is noted that, for skewed distributions, the bias of the estimators under imperfect ranking is the same as the bias under perfect ranking, as given in Subsection 4.1, because the expected value of the error term is zero.

To evaluate the performance of the MARSS design under ranking error, a simulation study was conducted. We generated 10,000 samples from a *N(0,1)* distribution, with errors generated from a normal distribution with variances $\sigma_e^2 = 0.1, 0.3$, and 0.5. The RP was simulated for different scenarios of MARSS at *h = 0, 1, 2* and *m = 3, 4, 5, 6, 7, 8*, using the equation:

$$RP_{MLRSS,Imp} = \frac{\frac{\sigma^2}{mc}}{V\left(\hat{\mu}_{MARSS,Imp}\right)} = \frac{\sigma^2}{mc}\Big/\frac{1}{10{,}000}\sum_{l=1}^{10{,}000}\left(\hat{\mu}^{(l)}_{MARSS,Im[p} - \mu\right)^2, \tag{23}$$

where $\hat{\mu}^{(l)}_{MARSS,Imp}$ is the mean of the l^{th} simulated sample.

Table 4 reports the RP of the MARSS estimators of *N(0,1)* under imperfect ranking for different levels of ranking error and set sizes. The results show that RP decreases as the ranking error increases, reflecting the impact of imperfect ranking on efficiency. Across all set sizes *m*, the RP of MARSS with *h = 1* and *h = 2* consistently outperforms RSS. For the normal distribution with ranking error, the results of RP further indicate that the MARSS estimator with *h = 2* is the most efficient (Figure 3).

Table 4. RP of MARSS mean estimators under imperfect ranking for $N(0,1)$.

	$k=0\ (RSS)$			$k=1$			$k=2$		
m	$\sigma_e^2=0.1$	$\sigma_e^2=0.3$	$\sigma_e^2=0.5$	$\sigma_e^2=0.1$	$\sigma_e^2=0.3$	$\sigma_e^2=0.5$	$\sigma_e^2=0.1$	$\sigma_e^2=0.3$	$\sigma_e^2=0.5$
3	1.72	1.57	1.49	**2.06**	1.71	1.60	**	**	**
4	2.05	1.77	1.62	**2.51**	1.95	1.74	**	**	**
5	2.40	1.99	1.73	2.80	2.17	1.90	**2.93**	2.19	1.93
6	2.61	2.12	1.86	3.07	2.37	2.00	**3.10**	2.36	2.08
7	2.88	2.30	1.91	3.42	2.43	2.08	**3.63**	2.57	2.05
8	3.14	2.35	2.02	3.63	2.60	2.11	**3.77**	2.55	2.12

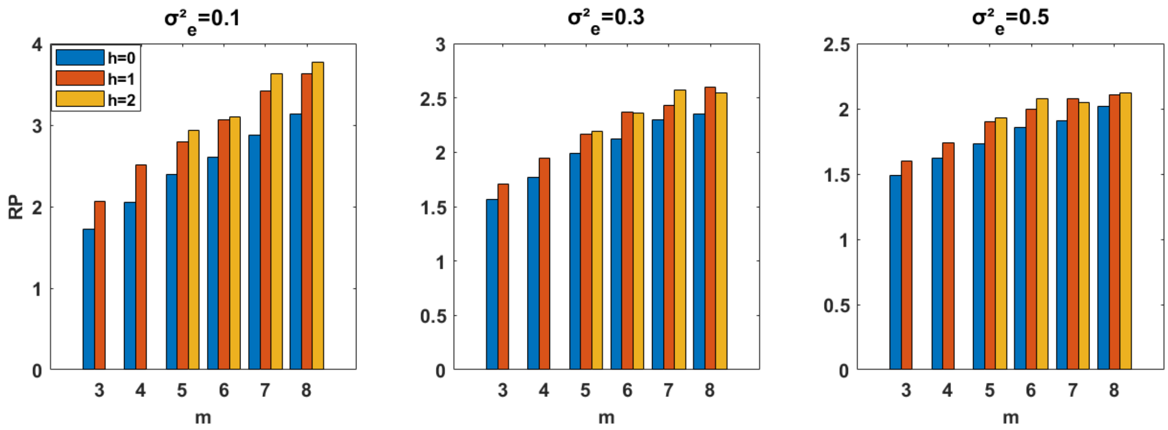

Figure 3. RP of MARSS and RSS mean estimators under imperfect ranking.

4.3. Concomitant variable effects

Since RSS and its variations rely on ranking units based on the researcher's judgment (either visually or using a concomitant variable), difficulties may arise when the variable of interest is hard to rank visually but is correlated with another variable. In such cases, using a concomitant variable becomes essential. In this subsection, we discuss the role of concomitant variables in ranking. The use of concomitant variables in RSS was first introduced by Stock [33], who considered the simple linear regression model of the variable of interest X on the concomitant variable Y:

$$X = \beta_0 + \beta_1 Y, \tag{24}$$

where the ranking is performed based on the variable Y, and the actual measurements are observed based on the corresponding X's. The regression model can be written as:

$$X = \mu_X + \rho \frac{\sigma_X}{\sigma_Y}(Y - \mu_Y), \tag{25}$$

where μ_X, μ_Y are the population means, σ_X, σ_Y are the standard deviations, and $\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$ is the correlation between X and Y. The estimator of μ under this model is

$$\hat{\mu}_{RSS,C} = \frac{1}{mc} \sum_{j=1}^{c} \sum_{i=1}^{m} X_{[i]}^{j}. \tag{26}$$

Using the laws of total expectation and variance, the mean and variance of $X_{[i]}$ are given by:

$$E\left(X_{[i]}\right) = E\left[E\left(X|Y_{(i)}\right)\right] = \mu_X + \rho\sigma_X E\left(\frac{Y_{(i)} - \mu_Y}{\sigma_Y}\right), \tag{27}$$

and

$$V\left(X_{[i]}\right)=E\left[V\left(X|Y_{(i)}\right)\right]+V\left[E\left(X|Y_{(i)}\right)\right]=\sigma_X^2(1-\rho^2)+\frac{\rho^2\sigma_X^2}{\sigma_Y^2}V\left(Y_{(i)}\right). \quad (28)$$

Hence, the variance of the RSS estimator is

$$V\left(\hat{\mu}_{RSS,C}\right)=\frac{1}{m^2c}\sum_{i=1}^{m}V\left(X_{[i]}\right). \quad (29)$$

The RP of the $\hat{\mu}_{RSS,C}$ compared to the SRS estimator is

$$RP_{RSS,C}=\frac{1}{1-\frac{\rho^2}{m}\sum_{i=1}^{m}\tau_{U_{(i)}}^2}, \quad (30)$$

where $U=\frac{Y-\mu_Y}{\sigma_Y}$ and $\tau_{U_{(i)}}=E\left(U_{(i)}\right)-E(U)$.

Based on the moment of $X_{[i]}$ in Eqs (26) and (27), the variance of the MARSS estimator $\hat{\mu}_{MLRSS,C}$ with a concomitant variable can be derived directly using the same moments employed in the design. The RP of the MARSS estimator compared to SRS is obtained as follows.

For odd *m*:

$$E\left(\hat{\mu}_{MARSS,Odd,C}\right)=\mu_X+\rho\frac{\sigma_X}{m\sigma_Y}\left[2h\left(\mu_{Y_{\left(\frac{m+1}{2}\right)}}-\mu_Y\right)+\left(\sum_{i=h+1}^{m-h}\mu_{Y_{(i)}}-(m-2h)\mu_Y\right)\right].$$

For even *m*:

$$E\left(\hat{\mu}_{MARSS,Even,C}\right)=\mu_X+\rho\frac{\sigma_X}{m\sigma_Y}\left[\left(h\mu_{Y_{\left(\frac{m}{2}\right)}}+h\mu_{Y_{\left(\frac{m}{2}+1\right)}}-2h\mu_Y\right)+\left(\sum_{i=h+1}^{m-h}\mu_{Y_{(i)}}-(m-2h)\mu_Y\right)\right].$$

The corresponding biases are:

$$B\left(\hat{\mu}_{MARSS,Odd,C}\right)=\rho\frac{\sigma_X}{m\sigma_Y}\left[2h\left(\mu_{Y_{\left(\frac{m+1}{2}\right)}}-\mu_Y\right)+\left(\sum_{i=h+1}^{m-h}\mu_{Y_{(i)}}-(m-2h)\mu_Y\right)\right] \quad (31)$$

and

$$B\left(\hat{\mu}_{MARSS,Even,C}\right)=\rho\frac{\sigma_X}{m\sigma_Y}\left[h\mu_{Y_{\left(\frac{m}{2}\right)}}+h\mu_{Y_{\left(\frac{m}{2}+1\right)}}-2h\mu_Y+\sum_{i=h+1}^{m-h}\mu_{Y_{(i)}}\right.$$
$$\left.-(m-2h)\mu_Y\right]. \quad (32)$$

The variances are:

$$V\left(\hat{\mu}_{MARSS,Odd,C}\right)=\frac{1}{m^2c}\left\{2h\sigma_{X_{\left[\frac{m+1}{2}\right]}}^2+\sum_{i=h+1}^{m-h}\sigma_{X_{[i]}}^2\right\}$$
$$=\frac{1}{m^2c}\left\{m\sigma_X^2(1-\rho^2)+\frac{\rho^2\sigma_X^2}{\sigma_Y^2}\left[2h\sigma_{Y_{\left(\frac{m+1}{2}\right)}}^2+\sum_{i=h+1}^{m-h}\sigma_{Y_{(i)}}^2\right]\right\}, \quad (33)$$

$$V\left(\hat{\mu}_{MARSS,Odd,C}\right)=\frac{1}{m^2c}\left\{2h\sigma_{X_{\left[\frac{m+1}{2}\right]}}^2+\sum_{i=h+1}^{m-h}\sigma_{X_{[i]}}^2\right\}$$
$$=\frac{1}{m^2c}\left\{m\sigma_X^2(1-\rho^2)+\frac{\rho^2\sigma_X^2}{\sigma_Y^2}\left[2h\sigma_{Y_{\left(\frac{m+1}{2}\right)}}^2+\sum_{i=h+1}^{m-h}\sigma_{Y_{(i)}}^2\right]\right\}. \quad (34)$$

Finally, the RP is defined as:

$$RP_{MARSS,C} = \frac{\sigma_X^2/mc}{MSE(\hat{\mu}_{MARSS,C})}. \tag{35}$$

To test the performance of the MARSS design using a concomitant variable *Y* for the target variable *X*, a simulation study was conducted with 10,000 samples generated from a bivariate normal distribution with mean 0 and variance 1. The ranking is conducted based on *Y,* while the actual measurement is taken on *X*. Different levels of correlation between the variable of interest *X* and its concomitant *Y* (ρ = *0.0, 0.25, 0.50, 0.75, 1.0*) were considered to evaluate the RP of MARSS (h = *0, 1, 2*) compared with the SRS estimator. Table 5 shows that the RP of the MARSS estimator increases with both correlation and set size (m). It is known that if $\rho = 0$, RSS and its modified designs become essentially random, so the RP is close to 1. At moderate ρ values, the RP of the MARSS estimator shows clear gains over SRS and RSS, while perfect correlation ($\rho = 1$) yields the highest efficiency, especially for larger m. The benefit of higher h values becomes evident as correlation strengthens, with MARSS at $h = 2$ performing best when the ranking is highly accurate (Figure 4).

Table 5. RP of the mean MARSS and RSS estimators using a concomitant variable.

m	$\rho = 0$ (SRS)	$\rho = 0.25$	$\rho = 0.50$	$\rho = 0.75$	$\rho = 1.00$
			$h = 0$		
3	0.98	1.03	1.15	1.38	1.96
4	0.99	1.06	1.17	1.46	2.37
5	1.00	1.07	1.16	1.54	2.77
6	1.03	1.02	1.19	1.65	3.17
7	1.01	1.05	1.21	1.73	3.57
8	1.00	1.04	1.25	1.77	4.07
			$h = 1$		
3	1.00	1.04	1.17	1.46	2.16
4	1.01	1.04	1.22	1.53	2.76
5	0.99	1.05	1.24	1.64	3.39
6	1.00	1.02	1.22	1.73	3.91
7	1.00	1.05	1.28	1.72	4.52
8	1.02	1.07	1.27	1.81	4.88
			$h = 2$		
5	0.98	1.05	1.21	1.70	3.48
6	1.01	1.02	1.28	1.78	4.05
7	0.98	1.06	1.23	1.81	4.69
8	1.00	1.05	1.28	1.86	5.35

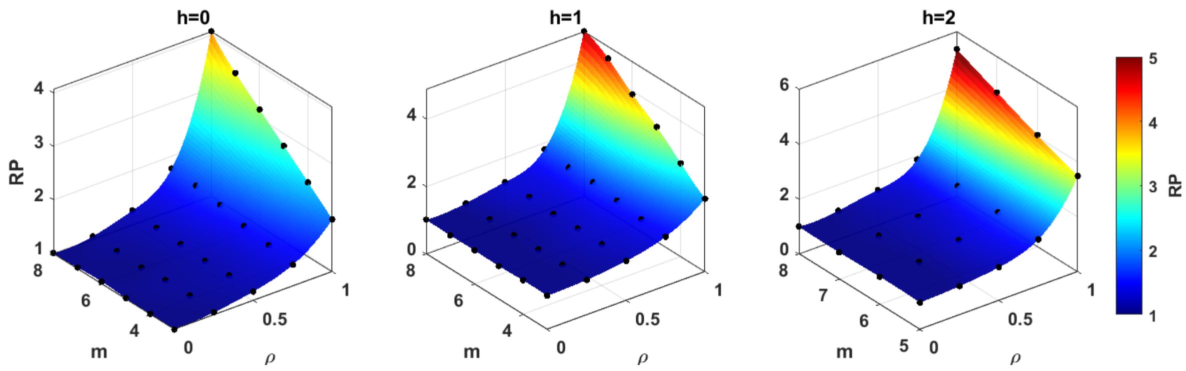

Figure 4. RP of MARSS and RSS estimators using a concomitant variable at different correlation levels.

4.4. Performance of outlier detection

To assess the performance of the MARSS design in handling outliers, we conducted a simulation study with 10,000 replications using SRS, RSS, and MARSS with $h = 1$ and $h = 2$. The simulations were carried out with cycles ($c = 10$) and various set sizes ($m = 3$ to 8). The methods were evaluated for *Normal(0,1)*. Outliers were injected into the simulated data at different contamination percentages (*CP = 5%, 10%, 15%,*) using the *Normal(10, 1)*.

To evaluate the robustness of each sampling method to the presence of outliers, we employed two performance measures. The first measure is the average percentage of outliers (PO) detected in the selected samples, calculated as:

$$PO = \frac{1}{10{,}000}\sum_{l=1}^{10{,}000}\left(\frac{\#Outliers_l}{n} \times 100\%\right), \tag{36}$$

where $n = mc$ is the sample size. Outliers were identified as any values outside the interval:

$$(Q_1 - 1.5 \times IQR, Q_3 + 1.5 \times IQR),$$

where Q_1 and Q_3 are the first and third quartiles, respectively, and $IQR = Q_3 - Q_1$ is the interquartile range. Lower PO values indicate greater robustness to contamination. The second measure is the RP of the mean estimators. It is simulated by:

$$RP = \frac{MSE(\hat{\mu}_{SRS})}{MSE(\hat{\mu}_{MARSS})}, \tag{37}$$

where $MSE(\hat{\mu}_{SRS})$ and $MSE(\hat{\mu}_{MARSS})$ represent the mean squared errors of the sample mean estimators obtained via SRS and MARSS, respectively. These MSEs were estimated using simulation under different sampling designs in the presence of outliers, and are computed as

$$MSE(\hat{\mu}_M) = \frac{1}{10{,}000}\sum_{l=1}^{10{,}000}\left(\hat{\mu}_M^{(l)} - \mu\right)^2, \tag{38}$$

where $\hat{\mu}_M^{(l)}$ is the sample mean from the l^{th} replication under the sampling design $M \in$ *{SRS, RSS (h = 0), MARSS(h = 1), MARSS (h = 2)}*.

Table 6 summarizes the average PO in samples drawn using SRS, RSS, and MARSS at different CP levels. The findings show that MARSS consistently produces fewer outliers than both RSS and SRS. Moreover, the proportion of outliers decreases as h increases (Figure 5). Overall, these results indicate that MARSS provides greater improvement and robustness in minimizing the influence of extreme values.

Table 7 and Figure 6 report the RP of the different estimators compared to SRS, in the presence of outliers. The results show that MARSS is generally more efficient than SRS and RSS in the presence of outliers. Overall, the findings highlight MARSS with *h* = *2* as a robust and highly efficient design, particularly effective in handling outliers.

Table 6. Outlier percentage in samples obtained via SRS, RSS, and MARSS from contaminated populations.

	CP = 5%				*CP = 10%*				*CP = 15%*			
m	SRS	*h = 0*	*h = 1*	*h = 2*	SRS	*h = 0*	*h = 1*	*h=2*	SRS	*h = 0*	*h = 1*	*h = 2*
3	5.98	5.96	2.51	**	10.39	10.48	4.34	**	13.38	14.13	7.21	**
4	5.84	5.80	2.27	**	10.28	10.50	4.21	**	13.52	14.37	6.97	**
5	5.77	5.71	2.24	1.69	10.44	10.45	3.76	2.55	14.52	14.95	6.12	4.22
6	5.66	5.63	2.13	1.54	10.35	10.35	3.87	2.47	14.49	15.07	6.50	4.08
7	5.69	5.62	2.43	1.56	10.35	10.35	4.14	2.31	14.91	15.22	6.71	3.65
8	5.63	5.57	2.30	1.57	10.35	10.27	4.22	2.37	14.76	15.13	7.08	3.75

Table 7. RP of mean estimators under MARSS at different outlier percentage.

		CP = 5%			*CP = 10%*			*CP = 15%*		
m	CP	*h = 0*	*h = 1*	*h = 2*	*h = 0*	*h = 1*	*h = 2*	*h = 0*	*h = 1*	*h = 2*
3	5	1.11	7.42	**	1.07	5.12	**	1.07	3.61	**
4	5	1.11	8.50	**	1.07	5.42	**	1.06	3.77	**
5	5	1.14	11.39	22.25	1.08	7.42	15.39	1.06	4.83	9.13
6	5	1.13	11.15	24.24	1.06	6.60	15.92	1.05	4.37	9.21
7	5	1.13	10.85	28.79	1.06	6.39	20.68	1.05	4.27	12.44
8	5	1.12	10.04	28.69	1.07	5.68	18.93	1.04	3.79	11.15

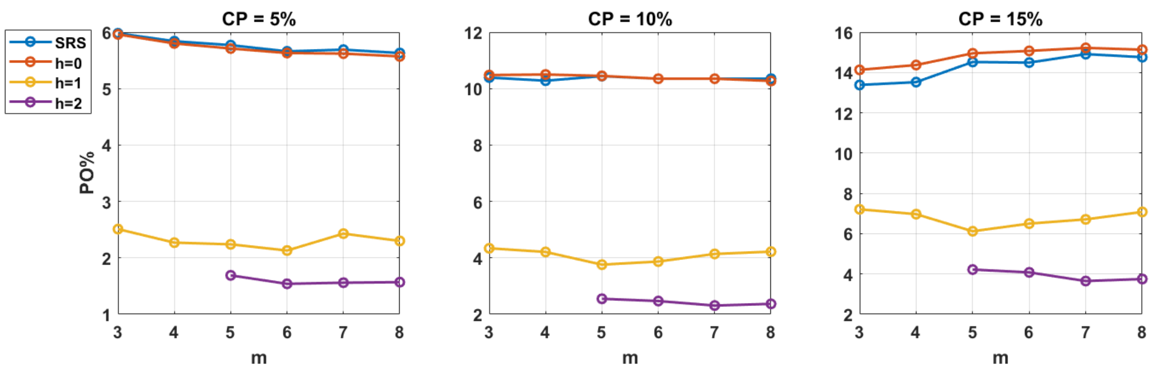

Figure 5. Outlier percentage in samples obtained via SRS, RSS, and MARSS.

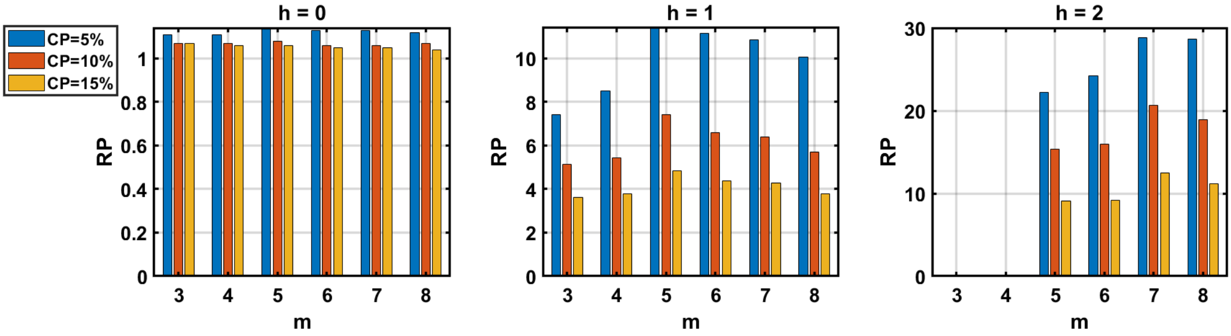

Figure 6. RP of the estimator with the presence of outliers.

4.5. Impact of ranking cost

The RSS design and its variations are often used to reduce data collection costs, especially when measurements are expensive but ranking is relatively easy. However, ranking itself may also involve costs. This issue was first addressed by Dell and Clutter [32], who evaluated the efficiency of RSS while accounting for ranking costs. In this framework, the total cost of data collection is expressed as the sum of measurement cost (C_m) and ranking cost (C_r). For SRS, only the measurement cost is considered. The cost relative efficiency (RE) compares the variance per cost unit between RSS and SRS. The adjusted-cost RE of the RSS estimator is expressed as:

$$RE_{cost} = RP.\frac{C_m}{C_m+C_r} = RP.\left(1+\frac{C_r}{C_m}\right)^{-1}. \tag{39}$$

Here, C_r/C_m represent the cost ratio of the ranking cost to the measurement cost. Generally, the use of RSS designs is more effective when measurements are expensive and ranking is inexpensive. However, this advantage decreases as the cost of ranking increases. From Eq (39), it can be denoted that the estimator becomes more efficient than SRS when RP is greater than $1+\frac{C_r}{C_m}$. Figure 7 presents the RE of the estimators for different RP values (*RP = 1.0, 1.2, 1.4, 1.6, 1.8, 2.0*). The results demonstrate that RE increases with either an increase in RP or a decrease in the cost ratio (C_r/C_m).

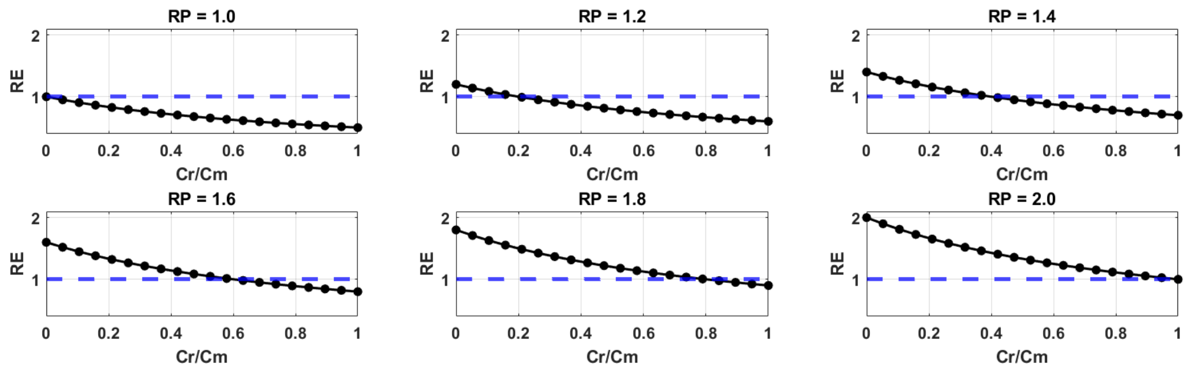

Figure 7. RE at different ranking cost ratios.

In summary, and based on the results, the MARSS method and its mean estimator offer the following:

- They help reduce measurement costs and are less affected by outliers than traditional RSS.

- They provide a smaller variance than both RSS and SRS when the distribution is unimodal and symmetric, and they remain unbiased for symmetric populations.
- They are more robust to outliers and keep good precision under different ranking conditions, including imperfect ranking or when using concomitant variables.
- They perform better than RSS in most cases, although they may show some bias when the distribution is skewed.

5. Compactness, Lindelöfness, and efficiency: a topological view of MARSS

In this section, we reinterpret the variance and robustness aspects in topological space. We model the population as a topological space (X, τ) with a probability law μ, and represent the ranking mechanism by a continuous function $f: X \to \mathbb{R}$. For a collection of open intervals $I_i \subset \mathbb{R}$, the preimages $U_i := f^{-1}(I_i)$ form the rank strips, which results in an open cover $\mathcal{U}_m = \{U_i\}_{i=1}^{m}$ of the support of μ. The MLRSS rule then appears as centralization in the first and last k strips together with diagonal selections in the remaining strips. Imperfect ranking is modeled by noisy scores $Y = f(X) + \varepsilon$, so that as $Var(\varepsilon)$ increases, the strip boundaries effectively thicken and the induced order weakens.

Lemma 6.1. Let Z be a real-valued random variable with a continuous distribution that is symmetric about its median Mz and unimodal. Then the variances of the order statistics decrease as one moves inward from the tails: for $i = 1, \dots, m$, the quantity $Var(Z_{(i:m)})$ is nonincreasing when i approaches the central index; for even m, the two middle order statistics achieve the minimal variance.

Theorem 6.2. Let (X, τ) be a topological space, $f: X \to \mathbb{R}$ continuous on the support of μ, and suppose $Z = f(X)$ is symmetric and unimodal about Mz. Fix $m \geq 3$ and $h \in \{0, 1, \dots, \lfloor (m-1)/2 \rfloor\}$. Consider one cycle of RSS and one cycle of MARSS constructed on the same m, h, and f. Then the MARSS sample mean has variance not exceeding that of RSS: $Var(\hat{\mu}_{MARSS}) \leq Var(\hat{\mu}_{RSS})$. If, in addition, Z is symmetric, then $E[\hat{\mu}_{MARSS}] = E[X] = \mu_X$.

Proof. Push the design through the ranking function *f*. The variance comparison becomes a weighted average of the ordered variances $Var(Z_{(i:m)})$. Relative to RSS, MLRSS replaces 2 h extreme-order contributions with 2 h central-order contributions, while leaving the diagonal contributions for $i = h+1, \dots, m-h$ unchanged. By Lemma 6.1 the ordered variances are smaller at the center than at the extremes, so the overall average cannot increase. Symmetry of Z implies that these central replacements do not shift the mean, and the diagonal selections average to the center, yielding unbiasedness.

Corollary 6.3. Let the observed ranking score be $Y = f(X) + \varepsilon$, where ε is independent noise with mean zero and finite variance. As $Var(\varepsilon) \to \infty$, the MLRSS selection becomes asymptotically uninformative and $RE = MSE(\hat{\mu}_{SRS}) / MSE(\hat{\mu}_{MARSS}) \to 1$.

Proposition 6.4. If $f(X)$ is Lindelöf, then for any $\varepsilon > 0$ there exists $m(\varepsilon)$ such that an MLRSS design using the first $m(\varepsilon)$ rank strips has mean squared error within ε of the ideal design based on the full countable cover induced by f.

Example 6.5. Let X be a cohort, f is the weight, and the measured variable is body-fat percentage. Empirically $f(X)$ lies in a compact interval. With $m = 5, h = 1$, define five rank strips by weight quantiles. MARSS picks the within-strip median in the outer two strips and the diagonal elements in the middle strip. The cover interpretation is $f^{-1}(I_i)$ with I_i as the weight intervals; centralization lowers the average ordered variance.

Example 6.6. Take $X = \mathbb{R}, f = id$, and dyadic intervals $I_j = (j2^{-r}, (j+1)2^{-r})$ inside $[-M, M]$ with $M \to \infty, r \to \infty$. By Proposition 6.4, choose M, r so that the truncated design with m strips has MSE within any prescribed ε of the limit. This yields a practical MARSS with theoretically guaranteed error control.

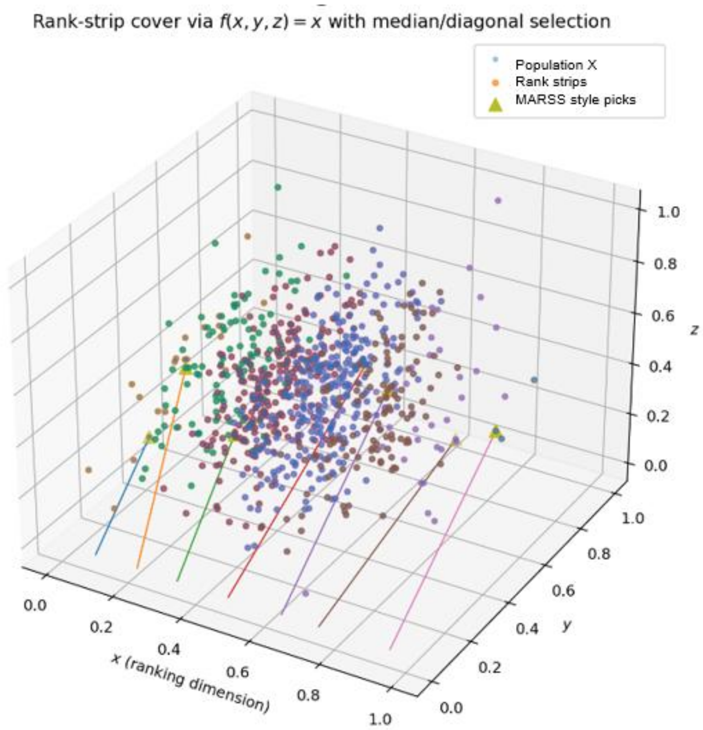

Figure 8. Visualization of an MARSS rank-strip design. The gray point cloud represents the population $x \subset \mathbb{R}^3$. Ranking is induced by $f(x, y, z) = x$, which partitions x into m rank strips (preimages of equal x-intervals). Triangular marker. Represents MARSS picks.

The topological interpretation of MARSS yields practical design rules. When $f(X)$ is compact, any finite interval cover of $f(X)$ pulls back via f to a finite rank-strip cover on X, so a finite-cycle design is feasible. If $f(X)$ is Lindelöf, one may work with a countable interval cover and truncate while controlling the mean squared error. In product populations $X = X_1 \times X_2$ with $f(x_1, x_2) = g_1(x_1) + g_2(x_2)$, the resulting rank strips are Minkowski-type sums of the one-dimensional strips; subspaces inherit strips by intersection. Choosing an auxiliary variable amounts to selecting a measurable f that is correlated with the response; larger correlation ρ sharpens strip separation and improves ordered-variance profiles. Increasing m refines the cover; the median-centric replacements stabilize the average of ordered variances, consistent with the observed efficiency gains under symmetry. Conversely, when the pushforward $f_{\#\mu}$ is skewed, centralization can introduce bias; in practice, $h = 1$ often trades a small bias for robust variance reduction.

6. Application to real data

The Body Fat Prediction Dataset [34], which includes $N = 252$ adult males, was used to test the applicability of the new sampling design, MARSS. This dataset records multiple characteristics, including body density, body fat percentage, age, weight, height, and measurements of several body parts such as neck, chest, abdomen, hips, thighs, knees, ankles, biceps, forearms, and wrists. These variables give a general picture of the health and physical characteristics of individuals [34]. Among them, body fat percentage (BFP) is a key health indicator, calculated using Siri's (1956) formula:

$$PBF = 495/density - 450.$$

In this application, the MARSS design was used to estimate the mean BFP. Measuring body is often more costly and complex than other variables, such as weight, which can be easily measured and ranked. Due to the strong correlation between weight and BFP, weight was used as a concomitant variable to assist in ranking the target variable (body fat percentage). Table 8 presents descriptive statistics for BFP (*X*) and weight (*Y*), including mean, dispersion measures, and skewness. The correlation coefficient (ρ) between the two variables is also reported. BFP has a mean (μ_X) of 19.15 with a standard deviation (σ_X) of 8.37, while weight shows higher variability ($\sigma_Y = 29.39$) around its mean ($\mu_Y = 178.92$). BFP is approximately symmetric (skewness ≈ 0.15), while the weight is right-skewed (skewness ≈ 1.20), indicating heavier tails in higher weight values (Figure 10). The positive correlation ($\rho = 0.61$) indicates a strong linear relationship between BFP and weight. Figure 10 further presents this relationship, showing a clear linear trend between the concomitant variable (weight) and the target variable (BFP). To illustrate the MARSS sampling design, a numerical example is provided, based on the scenario with $m = 5$, $c = 2$, and $h = 1$, with full implementation steps presented in Table 9.

To further evaluate the performance of the MARSS design in real-world data, we assessed the sample mean using several measures, including the expected value, bias, mean squared error (MSE), and RP. The simulation experiment was conducted using MATLAB R2024a, with 10,000 repetitions, and the performance measures were computed as follows:

$$E(\hat{\mu}) = \frac{1}{10{,}000}\sum_{l=1}^{10{,}000} \hat{\mu}_l\,, Bias(\hat{\mu}) = E(\hat{\mu}) - \mu_x, MSE(\hat{\mu}) = \frac{1}{10{,}000}\sum_{l=1}^{10{,}000} (\hat{\mu}_l - \mu_x)^2,$$

and

$$RP = \frac{MSE(\hat{\mu}_{SRS})}{MSE(\hat{\mu}_{MARSS})},$$

where $\hat{\mu}_l$ represents the estimate from the l^{th} replication, and μ_x denotes the true population mean.

Table 8. Descriptive statistics of body fat (*X*) and weight (*Y*).

	Mean (μ)	Std. Dev. (σ)	Variance (σ^2)	Skewness	Correlation (ρ)
Bodyfat (X)	19.1508	8.3687	70.0358	0.1455	0.61
Weight (Y)	178.9244	29.3892	863.7227	1.1981	

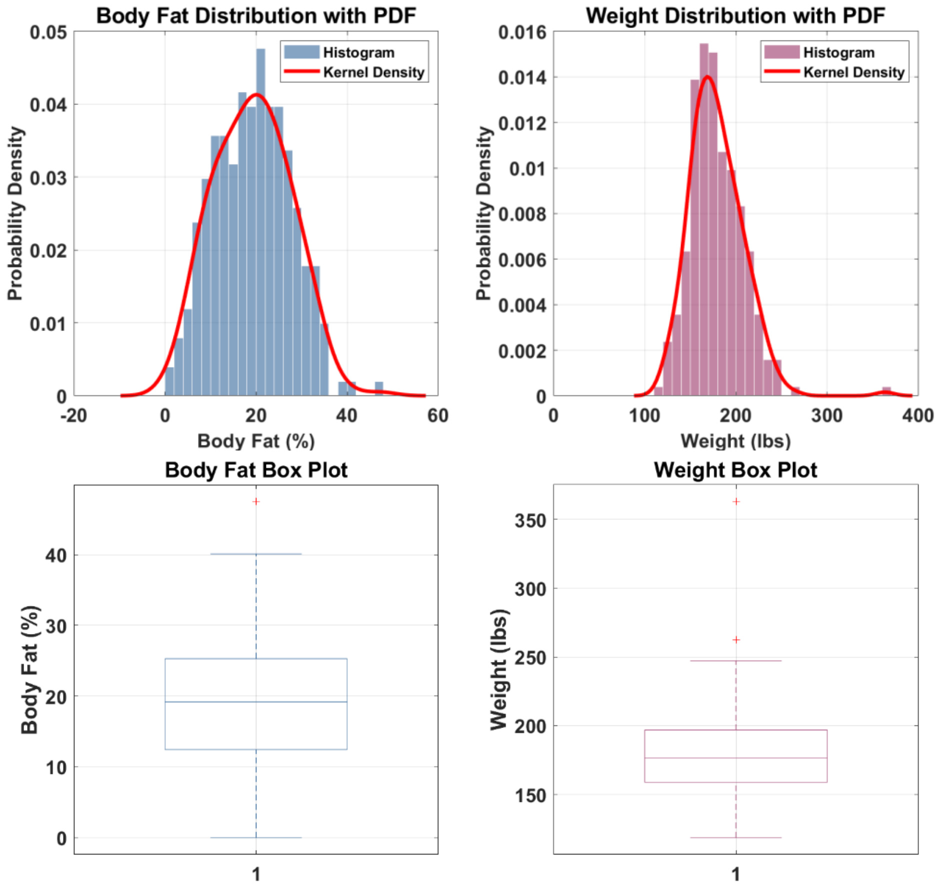

Figure 9. Distributions of weight and body fat.

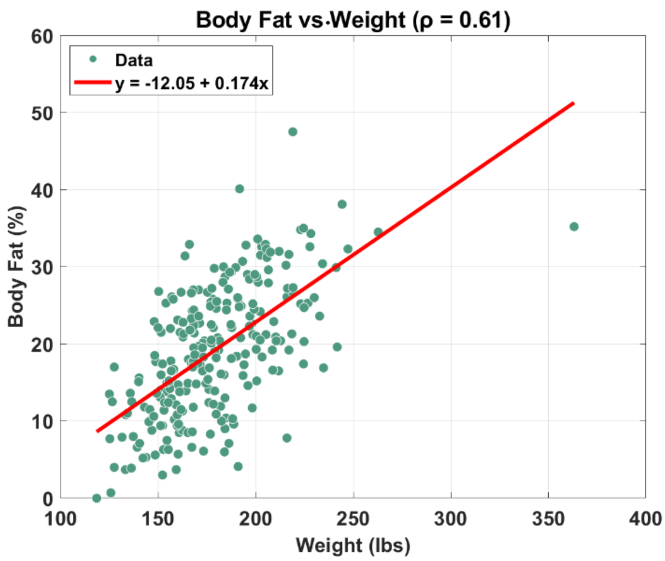

Figure 10. Scatter plot of weight versus BFP (%) with a fitted linear regression line.

Table 9. Numerical implementation of MARSS for body fat estimation (*m = 5, c = 1, h = 1*).

Set	Step1: Random indices	Step 2: Ranking based on weights	Step 3: Selected rank	Step 4: Measuring body fat
	Cycle 1			
1	3, 7, 12, 18, 22	154 < 181 < **200.5 (15.2)** < 209.25 < 216	3	15.2
2	5, 9, 14, 20, 25	151.25 < **184.25 (28.7) <** 191 < 205.25 < 211.75	2	28.7
3	2, 6, 11, 17, 24	148.75 < 173.25 < **186.25 (7.1) <** 195.75 < 210.25	3	7.1
4	4, 8, 13, 19, 23	140.25 < 176 < 180.5 < **183.75 (16.0) <** 184.75	4	16.0
5	1, 10, 15, 21, 16	154.25 < 162.75 < **179 (19.1) <** 187.75 < 198.25	3	19.1
	Cycle 2			
Set	Step1: Random indices	Step 2: Ranking based on weights	Step 3: Selected rank	Step 4: Measuring body fat
1	26, 34, 42, 58, 67	159.25 < 175.5 < **193.5 (15.9) <** 218.5 < 227.75	3	15.9
2	29, 37, 45, 53, 72	133.25 < **135.75 (13.6)** < 151 < 154.5 < 190.75	2	13.6
3	31, 39, 47, 61, 70	127.5 < 134.25 < **148.25 (18.5) <** 159.25 < 160.75	3	18.5
4	33, 41, 49, 63, 74	153 < 158.25 < 167 < **168 (11.8) <** 207.5	4	11.8
5	36, 44, 52, 65, 68	155.5 < 162.75 < **189.75 (29.9) <** 191.75 < 199.25	3	29.9
	$\hat{\mu}_{MLRSS} = (15.2 + 28.7 + 7.1 + 16.0 + 19.1 + 15.9 + 13.6 + 18.5 + 11.8 + 29.9)/10 = 18.19\%$			

Table 10 compares the performance of SRS, RSS (*h = 0),* and MARSS (*h = 1*) estimators for BFP across different sample sizes (*m = 3, 4, 5*) with 10 cycles (*c = 10*). The results indicate that SRS is nearly unbiased, while RSS maintains minimal bias and MARSS introduces a slight positive bias. Both RSS and MARSS outperform SRS in terms of MSE and RP (*RP > 1*), and MARSS has the highest efficiency for all cases.

Table 10. Performance measures of MARSS and RSS estimators for body fat.

m	Method	$E(\hat{\mu})$	B	MSE	RP
3	SRS	19.15	0.00	1.27	**
	RSS $(h = 0)$	19.13	-0.02	0.99	1.29
	MARSS $(h = 1)$	19.22	0.07	0.94	1.34
4	SRS	19.17	0.02	0.88	**
	RSS $(h = 0)$	19.15	0.00	0.60	1.51
	MARSS $(h = 1)$	19.21	0.06	0.57	1.54
5	SRS	19.16	0.01	1.13	**
	RSS $(h = 0)$	19.15	0.00	0.78	1.45
	MARSS $(h = 1)$	19.22	0.06	0.72	1.52

7. Conclusions

The study proposed a new sampling design for estimating the population mean. The new estimator is an unbiased estimator with lower variance than SRS and RSS in most scenarios. The design is effective when $h = 2$ for symmetrical unimodal distribution, and when $h = 1$ for the most skewed distribution, giving better efficiency and greater robustness to outliers. The estimator also performs well under imperfect ranking and when using a concomitant variable.

This approach can be extended to other estimation problems, such as estimating distribution parameters, reliability models, distribution functions, hazard rates, and variance and median. This approach can be used for different practical applications where reducing measurement cost is important.

Author contributions

Mahmoud Zuhier Aldrabseh: Conceptualization, methodology, formal analysis, simulation, writing original draft preparation, investigation, and visualization. Khudhayr A. Rashedi: Data curation, software implementation, validation, and critical review. Ali A. Atoom: Theoretical support and validation. Tariq S. Alshammari: Literature review, result interpretation, resources, and writing review and editing. Nisrein Al-Elaimat: Formatting and proofreading. All authors have read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

The authors declare that they have not used Artificial Intelligence (AI) tools in the creation of this article.

Data availability

The data used in this study are available from the corresponding author upon reasonable request.

Acknowledgments

This research has been funded by the Scientific Research Deanship at University of Ha'il, Saudi Arabia, through project number RG-24 067.

Conflict of interest

The authors declare that they have no conflicts of interest.

References

1. A. Haq, J. Brown, E. Moltchanova, A. I. Al-Omari, Partial ranked set sampling design, *Environmetrics*, **24** (2013), 201–207. https://doi.org/10.1002/env.2203
2. G. A. McIntyre, A method for unbiased selective sampling using ranked sets, *Aust. J. Agr. Res.*, **3** (1952), 385–390. https://doi.org/10.1071/AR9520385
3. H. A. Muttlak, Median ranked set sampling, *J. Appl. Stat. Sci.*, **6** (1997), 245–255.
4. M. T. Al-Odat, M. F. Al-Saleh, A variation of ranked set sampling, *J. Appl. Stat. Sci.*, **10** (2001), 137–146.
5. A. D. Al-Nasser, L ranked set sampling: a generalization procedure for robust visual sampling, *Commun. Stat. Simul. Comput.*, **36** (2007), 33–43. https://doi.org/10.1080/03610910601096510

6. L. Khan, J. Shabbir, U. Khalil, A new systematic ranked set-sampling scheme for symmetric distributions, *Life Cycle Reliab. Saf. Eng.*, **8** (2019), 205–210. https://doi.org/10.1007/s41872-019-00080-5
7. M. Z. Aldrabseh, M. T. Ismail, New modification of ranked set sampling for estimating population mean, *J. Stat. Comput. Simul.*, **93** (2023), 2843–2855. https://doi.org/10.1080/00949655.2023.2212312
8. M. F. Al-Saleh, M. A. Al-Kadiri, Double-ranked set sampling, *Stat. Probab. Lett.*, **48** (2000), 205–212. https://doi.org/10.1016/S0167-7152(99)00206-0
9. M. F. Al-Saleh, A. I. Al-Omari, Multistage ranked set sampling, *J. Stat. Plan. Infer.*, **102** (2002), 273–286. https://doi.org/10.1016/S0378-3758(01)00086-6
10. A. I. Al-Omari, A. Haq, A new sampling method for estimating the population mean, *J. Stat. Comput. Simul.*, **89** (2019), 1973–1985. https://doi.org/10.1080/00949655.2019.1604710
11. M. Z. Aldrabseh, M. T. Ismail, A. I. Al-Omari, Double except extreme ranked set sampling for estimating population mean, *Adv. Math. Model. Appl.*, **9** (2024), 415–430. https://doi.org/10.62476/amma93415
12. M. Noor-ul-Amin, F. Arif, M. Hanif, Modified extreme ranked sets sampling with auxiliary variable, *Proc. Natl. Acad. Sci., India, Sect. A Phys. Sci.*, **91** (2021), 537–542. https://doi.org/10.1007/s40010-020-00698-6
13. M. A. Shehzada, A. Nisara, A. Khana, W. Emamb, Y. Tashkandyb, H. Khurram, et al., Modified median quartile double ranked set sampling for estimation of population mean, *Heliyon*, **10** (2024), e34627. https://doi.org/10.1016/j.heliyon.2024.e34627
14. A. Haq, J. Brown, E. Moltchanova, A. I. Al-Omari, Paired double-ranked set sampling, *Commun. Stat. Theory Methods*, **45** (2016), 2873–2889. https://doi.org/10.1080/03610926.2014.892135
15. A. I. Al-Omari, M. S. Abdallah, Estimation of the distribution function using moving extreme and MiniMax ranked set sampling, *Commun. Stat. Simul. Comput.*, **52** (2023), 1909–1925. https://doi.org/10.1080/03610918.2021.1891433
16. A. I. Al-Omari, M. Z. Raqab, Estimation of the population mean and median using truncation-based ranked set samples, *J. Stat. Comput. Simul.*, **83** (2013), 1453–1471. https://doi.org/10.1080/00949655.2012.662684
17. S. Asghari, B. S. Gildeh, J. Ahmadi, G. M. Borzadaran, Sign control chart based on ranked set sampling, *Qual. Technol. Quant. Manag.*, **15** (2018), 568–588. https://doi.org/10.1080/16843703.2017.1314094
18. A. Mohammadkhani, A. Amiri, M. B. C. Khoo, A review of ranked set sampling and modified methods in designing control charts, *Qual. Reliab. Eng. Int.*, **39** (2023), 1465–1493. https://doi.org/10.1002/qre.3282
19. M. Z. Aldrabseh, M. T. Ismail, A. I. Al-Omari, Except-extremes ranked set sampling for estimating the population variance with two applications of real data sets, *Aust. J. Stat.*, **53** (2024), 99–113. https://doi.org/10.17713/ajs.v53i4.1872
20. P. L. H. Yu, K. Lam, Regression estimator in ranked set sampling, *Biometrics*, **53** (1997), 1070–1080. https://doi.org/10.2307/2533564
21. X. Dong, L. Zhang, Estimation of system reliability for exponential distributions based on L ranked set sampling, *Commun. Stat. Theory Methods*, **49** (2020), 3650–3662. https://doi.org/10.1080/03610926.2019.1691735
22. S. A. Benchiha, A. I. Al-Omari, G. Alomani, Enhanced estimation of the unit Lindley distribution parameter via ranked set sampling with real-data application, *Mathematics*, **13** (2025), 1645. https://doi.org/10.3390/math13101645

23. M. Z. Aldrabseh, M. T. Ismail, A. I. Al-Omari, Enhanced estimation of strength–stress reliability using except extreme ranked set sampling under exponential distribution, *Life Cycle Reliab. Saf. Eng.*, 2025. https://doi.org/10.1007/s41872-025-00326-5
24. M. Chen, W. X. Chen, R. Yang, Y. W. Zhou, Exponential–Poisson parameters estimation in moving extremes ranked set sampling design, *Acta Math. Appl. Sin. Engl. Ser.*, **41** (2025), 973–984. https://doi.org/10.1007/s10255-023-1076-1
25. A. Kumar, S. Bhushan, R. Pokhrel, Small area estimation under ranked set sampling: simulation and real-life application with crop production data, *Qual. Quant.*, 2025, 1–21. https://doi.org/10.1007/s11135-025-02422-1
26. A. Kumar, S. Bhushan, R. Pokhrel, Direct and synthetic logarithmic estimators in small area estimation via ranked set sampling, *Statistics*, 2025, 1–22. https://doi.org/10.1080/02331888.2025.2542350
27. A. S. Hassan, D. S. Metwally, H. E. Semary, S. A. Benchiha, A. M. Gemeay, M. Elgarhy, Improved estimation based on ranked set sampling for the Chris–Jerry distribution with application to engineering data, *Comput. J. Math. Stat. Sci.*, **4** (2025), 424–456. https://doi.org/10.21608/cjmss.2025.375962.1156
28. G. Alomani, S. A. Benchiha, A. I. Al-Omari, Estimation of the inverse power Lindley distribution parameters using ranked set sampling with an application to failure time data, *Axioms*, **14** (2025), 801. https://doi.org/10.3390/axioms14110801
29. H. Kittani, M. Alaesa, G. Gharib, Comparison among some methods for estimating the parameters of truncated normal distribution, *J. Adv. Math.*, **20** (2021), 79–95. https://doi.org/10.24297/jam.v20i.8934
30. K. Takahasi, K. Wakimoto, On unbiased estimates of the population mean based on the sample stratified by means of ordering, *Ann. Inst. Stat. Math.*, **20** (1968), 1–31. https://doi.org/10.1007/BF02911622
31. T. R. Dell, J. L. Clutter, Ranked set sampling theory with order statistics background, *Biometrics*, **28** (1972), 545–555. https://doi.org/10.2307/2556166
32. H. A. David, H. N. Nagaraja, *Order statistics*, 3 Eds., Wiley Series in Probability and Statistics, John Wiley & Sons, Hoboken, NJ, 2003. https://doi.org/10.1002/0471722162
33. S. L. Stokes, Ranked set sampling with concomitant variables, *Commun. Stat. Theory Methods*, **6** (1977), 1207–1211. https://doi.org/10.1080/03610927708827563
34. Kaggle, Casper6290. BodyFat Prediction [Data set]. Kaggle. Available from: https://www.kaggle.com/code/casper6290/bodyfat-prediction.

AIMS Press

©2025 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0)